DOCUMENT RESUME

ED 343 914 TM 018 045

AUTHOR	Tucker, Mary L.; Daniel, Larry G., Jr.
TITLE	Investigating Result Stability of Canonical Function Equations with the Jackknife Technique.
PUB DATE	Jan 92
NOTE	12p.; Paper presented at the Annual Meeting of the Southwest Educational Research Association (Houston, TX, January 31-February 2, 1992).
PUB TYPE	Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	College Faculty; *Correlation; *Equations (Mathematics); *Estimation (Mathematics); Mathematical Models; *Multivariate Analysis; *Reliability; Sample Size
IDENTIFIERS	*Invariance; *Jackknifing Technique; Research Replication

ABSTRACT

The jackknife statistic is discussed as a viable invariance procedure. Data from a study of leadership illustrates the use of the jackknife in determining the stability of canonical function coefficients following canonical correlation analysis. The jackknife procedure entails arbitrarily omitting one observation or a subset of observations at a time from the original sample and recalculating the original statistical estimator for each of the resulting truncated data sets. The procedure is repeated with each individual observation or unique subgroup omitted. Pseudovalues are computed for each of the truncated data sets, based on the computation of the original and the sample-minus-one subset function coefficients. These pseudovalues are then averaged, providing a jackknifed estimate of the canonical function coefficients. The stability of the original values is gauged by determining whether they fall within confidence intervals for the jackknifed values. The sample illustrating the jackknife statistic is taken from a study by M. L. Tucker of leadership styles with data from 106 college faculty and administrators who rated their supervisors. Because the jackknife technique minimizes sample splitting through sample omission and reuse, it is particularly useful when the sample size is small. There is a 10-item list of references and one table with analysis results. (SLD)

ED343914

INVESTIGATING RESULT STABILITY OF CANONICAL FUNCTION EQUATIONS WITH THE JACKKNIFE TECHNIQUE

Mary L. Tucker, Ph.D.
College of Business
Nicholls State University

Larry G. Daniel, Jr., Ph.D.
Education Leadership and Research
The University of Southern Mississippi

PLEASE CORRESPOND WITH:

Dr. Mary L. Tucker, Assistant Professor
Office Information Systems Department
College of Business, POB 2042
Nicholls State University
Thibodaux, LA 70310
(504) 448-4206 / FAX (504) 448-4922

Paper presented at the annual meeting of the Southwest Educational Research Association, Houston, TX, January 28, 1992.

ABSTRACT

The jackknife statistic is discussed as a viable invariance procedure. Data from a substantive leadership study is used to illustrate the value of the jackknife statistic in determining the stability of canonical function coefficients following a canonical correlation analysis.

Jackknife Statistic

The jackknife statistic was refined by Tukey (1958) from research reported by Quenouille (1949, 1956) and Jones (1956) to provide a measure of stability and significance of results. The essentia of the jackknife approach, according to Crask and Perreault (1977), "is to partition out the impact of effect of a particular subset of data on an estimate derived from the total sample" (p. 61).

Fenwick (1979) described the jackknife method as versatile, offering "the researcher the opportunity to reduce bias, perform significance tests, and assess the validity and stability of analyses without requiring a large sample" (p. 410). The jackknife statistic, then, is useful with different statistical procedures and is especially appropriate when sample size is small, as was the case in the present study where multivariate procedures were utilized.

Overview of the Jackknife Statistic Computation

The jackknife procedure entails arbitrarily omitting one observation or subset of observations at a time from the original sample and recalculating for each of the resulting truncated data sets the orginal statistical estimator. This procedure is repeated, with each individual observation or unique subgroup, in turn omitted. "Pseudovalues" (Quenouille, 1956) are computed for each of the truncated data sets, based upon the computation of the original and the sample-minus-one subset function coefficients. These pseudovalues are then averaged, providing a "jackknifed" estimate of the canonical function coefficients. Stability of the original values is gauged by determining whether they fall within confidence intervals for the jackknifed values (Daniel, 1989).

Daniel (1989, p. 7) provides the following documentation for compilation of the jackknife statistic.

> A given sample of size *N* is partitioned into *k* subsets of size *M* (*kM* = *N*). All subsets must be of the same size (*M*) and may be as small as one case or as large as the largest multiplicative factor of *N*. A predictive estimator (e.g., a discriminant function [or canonical function] coefficient), designated as theta-prime (0') is then computed using all *k* of the subsamples from the original sample of size *N*. The same estimator is also computed with the i^{th} subset (i = 1 to *k*) omitted from the sample. This estimator is designated as 0_i'. This procedure is repeated *k* times with a different subset omitted each time. Before computing the jackknifed estimator, weighted combinations of the 0' and 0_i' values are computed. These weighted values are called pseudovalues ... and are designated by the letter *J*. The pseudovalues are computed using the equation:
>
> (1) J_i (0') = k 0' - (k - 1) 0_i'
>
> where i = 1, 2, 3, ..., *k*.
>
> The average of the pseudovalues is the jackknifed estimator:
>
> (2) J (0') = [SUM J_i (0')] / k
>
> where i = 1, 2, 3, ..., *k*.

Tukey (1958) postulated that a given set of pseudovalues could be regarded as an approximately normal distribution; hence the stability of a given jackknifed estimator may be evaluated by determining confidence intervals about the estimator, and then testing to determine whether the researcher can conclude that the population estimator falls within those confidence interval bands.

From this analysis, the researcher can ascertain fluctuations in sampling error that may occur because of a single deleted observation's uniqueness or because of the combined characteristics of cases within a given subset (Campo, 1988; Davidson & Giroir, 1989; Diaconis & Bradley, 1983). Crask & Perreault (1977) provide an illustrative example of the jackknife application, with a brief non-technical explanation of the mathematical procedures involved in computation of the jackknife statistic.

Application of the Jackknife Technique

The sample used for illustration of the jackknife statistic is taken from a substantive study (Tucker, 1990) of transformational, transactional, and laissez-faire leadership styles as predictors of follower satisfaction with leader, of follower perceptions of leader effectiveness, and of extra effort expended by the follower for a leader. Data were gathered from 106 faculty and administrative respondents who rated their individual leaders located at a Southern, urban university (XYZ).

The XYZ data ($n = 106$) in hand were partitioned into 53 subsets of two cases each. Predictive estimators (standardized canonical function coefficients) were calculated using the entire original sample size ($n = 106$) and then were calculated for each of the partitioned subsamples ($n = 104$). The averaged weighted value of the estimator when the analysis is run repeatedly with the various subsets of data was used to compute the jackknifed value of the estimator. The analysis yielded three canonical functions. Standardized canonical function coefficients derived from the entire data set for the three predictor variables for Functions I, II, and

III were, respectively, .983, .010, and .026 (Transformational Leadership); .912, 1.33, and .288 (Transactional Leadership); and .61, .19, and 1.15 (laissez-faire). Estimators were then computed 53 additional times for Functions I, II, and III dropping one different subsample from the canonical correlation analysis each time.

Weighted pseudovalues were computed separately for each of the three canonical function coefficients (estimators) over each of the 53 subsequent analyses using equation (1). Pseudovalues derived from these estimators were averaged for the jackknife estimator. Tukey (1958) reasoned that a given set of pseudovalues could be theorized as an approximation of a normal distribution; therefore, the stability of a jackknifed estimator may be judged by determining confidence intervals about the estimator. Determination can then be made as to whether the population estimator falls within those confidence interval bands. This test is administered by dividing the estimator by its associated standard error to obtain a t-value. A jackknifed estimator is considered stable, then, if its calculated t-value exceeds the t-critical value (Daniel, 1989). Jackknifed statistic standardized canonical function coefficients, and associated jackknifed coefficients with standard errors for each predictor variable are presented in Table 1. Confidence intervals of 95 percent for the jackknifed coefficients are also presented. Based on the standard error confidence intervals, the stability of the jackknifed coefficients for two of the three predictors were supported across three canonical functions. Transformational leadership (TF) and transactional leadership (TA) standardized canonical coefficients were within the confidence level; laissez-faire leadership (LF) coefficients were not.

Insert Table 1 About Here

Discussion

As noted previously, the jackknife statistic is useful in the stability evaluation of a given estimator by eliminating bias due to the inclusion of atypical or outlying observations in a particular sampling. In this research example, the stability of jackknifed canonical function coefficients for three predictors of respondents' perceived satisfaction, effectiveness, and extra effort was analyzed. The jackknifed coefficients for all three of the predictors were quite close in value to the original coefficients obtained using the entire sampling. Standard error confidence intervals and t-values were computed for each coefficient. Based upon these computations, the stability of the jackknifed coefficients for two of the three predictors (transformational leadership and transactional leadership) was supported while the stability of the coefficient for the third predictor (laissez-faire leadership) was not. These findings indicate that the first two predictors may be considered as valid discriminators between the outcome variables of satisfaction, effectiveness, and extra effort, and that the results may be appropriately generalized to the larger population of interest. As indicated by the data presented in Table 1, the third variable (laissez-faire leadership) tends to be unstable against changes in the composition of the sample, and therefore is a more biased indicator.

Summary

The jackknife invariance procedure has been promoted as superior to other traditional validation methods because it make use of all of the data in a particular data set while eliminating bias in estimates of stability by "averaging out" the effects of outlying or atypical observations within a given data set. Crask & Perreault (1977) illustrate that the use of these techniques produce less biased and more conservative estimates of true population characteristics. Because the jackknife technique minimizes sample splitting through sample omission and reuse, it is particularly useful when sample size is small (Fenwick, 1987). The foregoing analysis demonstrated an appropriate use of the jackknife statistic as a tool for assessing the stability of canonical discriminant function co -ficients following a canonical analysis procedure.

REFERENCES

Campo, S. (1988, November). Alternative logics for estimating whether research will generalize. Paper presented at the annual meeting of the Mid-South Educational Research Association, Louisville, KY.

Crask, M. R. & Perreault, W. D. (1977). Validation of discriminant analysis in marketing research. Journal of Marketing Research, 14, 60-67.

Daniel, L. G. (1989, January 27). Use of the jackknife statistic to establish the external validity of discriminant analysis results. Paper presented at the annual meeting of the Southwest Educational Research Association, Houston, TX. (ERIC Document Reporduction Services No. Ed 305 382)

Davidson, B. M. & Giroir, M. M. (1989, November 9). Ways of estimating the probability that results will replicate. Paper presented at the annual meeting of the Mid-South Educational Research Association, Little Rock, AR.

Diaconis, P., & Bradley, E. (1983). Computer intensive methods in statistics. Scientific American, 248 (5), 116-130.

Fenwick, I. (1979). Techniques in market measurement: The Jackknife. Journal of Marketing Research. XVI, 410-414.

Jones, H. L. (1956, March). Investigating the properties of a sample mean by employing random subsample means. Journal of the American Statistical Association, 51, 54-83.

Quenouille, M. (1949). Approximate tests of correlations in time series. Journal of Royal Statistical Social Series B, 11, 68-84.

Quenouille, M. (1956). Notes of bias in estimation. Biometrika, 43, 647-649.

Tucker, M. L. (1990). _Higher education leadership: transformational leadership as a predictor of satisfaction, effectiveness, and extra effort_. Unpublished doctoral dissertation, University of New Orleans, LA.

Table 1

Jackknife Statistic Coefficients, Standard Error, and Confidence Intervals

	Function I			Function II			Function III		
	TF	TA	LF	TF	TA	LF	TF	TA	LF
Std. Coef.	-.98	.91	.61	-.01	-1.33	-.19	.03	.29	1.15
Jck. Coef.	-.98	.91	.74	-.11	-1.37	-.35	.03	.31	1.26
Std. Error	.04	.04	.01	.00	.05	.01	.01	.02	.01
Confidence Intervals:									
Upper	-1.02	.95	.75	-.01	-1.32	-.34	.04	.33	1.27
Lower	- .94	.87	.73	-.01	-1.42	-.37	.02	.29	1.25
	*	*		*	*		*	*	

Note. Confidence Interval (.05, *df*=52) = Jck. coef. ± Std. error.
*Standardized coefficient falls within the confidence intervals.